

#2

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/911,513

DATE: 08/01/2001
 TIME: 17:10:22

Input Set : A:\Sequence Listing.txt
 Output Set: N:\CRF3\08012001\I911513.raw

ENTERED

4 <110> APPLICANT: Harberd, Nicholas P
 5 Peng, Jinrong
 6 Carol, Pierre
 7 Richards, Donald E
 9 <120> TITLE OF INVENTION: Nucleic acid encoding GAI gene of Arabidopsis thaliana
 11 <130> FILE REFERENCE: 620-157
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/911,513
 14 <141> CURRENT FILING DATE: 2001-07-25
 16 <150> PRIOR APPLICATION NUMBER: US 09/117,853
 17 <151> PRIOR FILING DATE: 1998-08-12
 19 <150> PRIOR APPLICATION NUMBER: PCT/GB97/00390
 20 <151> PRIOR FILING DATE: 1997-02-12
 22 <150> PRIOR APPLICATION NUMBER: GB 9602796.6
 23 <151> PRIOR FILING DATE: 1996-02-12
 25 <160> NUMBER OF SEQ ID NOS: 12
 27 <170> SOFTWARE: PatentIn Ver. 2.0
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 1964
 31 <212> TYPE: DNA
 32 <213> ORGANISM: Arabidopsis thaliana
 34 <400> SEQUENCE: 1
 35 taataatcat tttttttctt ataaccttcc tctctatattt tacaatttat tttgttatta 60
 36 gaagtggtag tggagtga aaacaaatcc taagcagtc taaccgatcc ccgaagctaa 120
 37 agattcttca ccttcccaaa taaagcaaaa ctagatccg acattgaagg aaaaaccttt 180
 38 tagatccatc tctgaaaaaa aaccaaccat gaagagagat catcatcatc atcatcaaga 240
 39 taagaagact atgatgatga atgaagaaga cgacggtaac ggcatggatg agcttctagc 300
 40 tgttcttggg tacaagggtta ggcatcgcga aatggctgat gttgctcaga aactcgagca 360
 41 gcttgaagtt atgatgtcta atgttcaaga agacgatctt tctcaactcg ctactgagac 420
 42 tgttcactat aatccggcgg agctttacac gtggcttgat tctatgctca ccgaccttaa 480
 43 tctccgctcg tctaacgccg agtacgatct taaagctatt cccggtgacg cgattctcaa 540
 44 tcagttcgct atcgattcgg cttcttcgtc taaccaaggc ggcgaggag atactgtatac 600
 45 tacaacaag cggttgaat gctcaaaccg cgtcgtgga accaccacag cgacggctga 660
 46 gtcaactcgg catgttgacc tggttgact gcaggagaa ggtgtgcgtc tegtacgcgc 720
 47 gcttttggct tgcgctgaag ctgttcagaa ggagaatctg actgtggcgg aagctctggt 780
 48 gaagcaaata ggattcttag ctgtttctca aatcgagct atgagaaaag tgcactacta 840
 49 ctctcgccga gctctcgcgc ggcgatttta ccgtctctct ccgtcgcaga gtccaatoga 900
 50 ccaactctct tccgatactc ttcagatgca cttctacgag acttgtcctt atctcaagtt 960
 51 cgctcacttc acggcgaatc aagcgattct cgaagctttt caagggaaga aaagagttca 1020
 52 tgtcattgat ttctctatga gtcaaggctt tcaatggcgg gcgcttatgc aggctcttgc 1080
 53 gcttcgacct ggtggctctc ctgttttccg gtttaaccgga attggtccac cggcacccgga 1140
 54 taatttcgat tatcttcatg aagttgggtg taagctggct catttagctg aggcgattca 1200
 55 cgttgagttt gagtacagag gatttgtggc taacacttta gctgatcttg atgcttcgat 1260
 56 gcttgagctt agaccaagt agattgaatc tgttgccggt aactctgttt togagcttca 1320
 57 caagctcttg ggacgacctg gtgcgatoga taaggttctt ggtgtggtga atcagattaa 1380
 58 accggagatt ttcactgtgg ttgagcagga atcgaaacct aatagtcoga ttttcttaga 1440
 59 tcggtttact gagtcgttgc attattactc gacgttggtt gactcgttgg aaggtgtacc 1500
 60 gagtgggtcaa gacaagggtc tgcggaggtt ttacttggtt aaacagatct gcaacgttgt 1560

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/911,513

TIME: 17:10:22

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\08012001\I911513.raw

```

61 ggcttgtgat ggacctgacc gagttgagcg tcatgaaacg ttgagtcagt ggaggaaccg 1620
62 gttcgggtct gctgggtttg cggctgcaca tattggttcg aatgcgttta agcaagcgag 1680
63 tatgcttttg gctctgttca acggcggtga gggttatcgg gtggaggaga gtgacggctg 1740
64 tctcatgttg ggttggcaca cacgaccgct catagccacc tcggcttga aactctccac 1800
65 caattagatg gtggctcaat gaattgatct gttgaaccgg ttatgatgat agatttccga 1860
66 ccgaagccaa actaaatcct actgtttttc cctttgtcac ttgttaagat cttatctttc 1920
67 attatattag gtaattgaaa aatttctaaa ttactcacac tggc 1964
69 <210> SEQ ID NO: 2
70 <211> LENGTH: 532
71 <212> TYPE: PRT
72 <213> ORGANISM: Arabidopsis thaliana
74 <400> SEQUENCE: 2
75 Met Lys Arg Asp His His His His His Gln Asp Lys Lys Thr Met Met
76 1 5 10 15
78 Met Asn Glu Glu Asp Asp Gly Asn Gly Met Asp Glu Leu Leu Ala Val
79 20 25 30
81 Leu Gly Tyr Lys Val Arg Ser Ser Glu Met Ala Asp Val Ala Gln Lys
82 35 40 45
84 Leu Glu Gln Leu Glu Val Met Met Ser Asn Val Gln Glu Asp Asp Leu
85 50 55 60
87 Ser Gln Leu Ala Thr Glu Thr Val His Tyr Asn Pro Ala Glu Leu Tyr
88 65 70 75 80
90 Thr Trp Leu Asp Ser Met Leu Thr Asp Leu Asn Pro Pro Ser Ser Asn
91 85 90 95
93 Ala Glu Tyr Asp Leu Lys Ala Ile Pro Gly Asp Ala Ile Leu Asn Gln
94 100 105 110
96 Phe Ala Ile Asp Ser Ala Ser Ser Ser Asn Gln Gly Gly Gly Gly Asp
97 115 120 125
99 Thr Tyr Thr Thr Asn Lys Arg Leu Lys Cys Ser Asn Gly Val Val Glu
100 130 135 140
102 Thr Thr Thr Ala Thr Ala Glu Ser Thr Arg His Val Val Leu Val Asp
103 145 150 155 160
105 Ser Gln Glu Asn Gly Val Arg Leu Val His Ala Leu Leu Ala Cys Ala
106 165 170 175
108 Glu Ala Val Gln Lys Glu Asn Leu Thr Val Ala Glu Ala Leu Val Lys
109 180 185 190
111 Gln Ile Gly Phe Leu Ala Val Ser Gln Ile Gly Ala Met Arg Lys Val
112 195 200 205
114 Ala Thr Tyr Phe Ala Glu Ala Leu Ala Arg Arg Ile Tyr Arg Leu Ser
115 210 215 220
117 Pro Ser Gln Ser Pro Ile Asp His Ser Leu Ser Asp Thr Leu Gln Met
118 225 230 235 240
120 His Phe Tyr Glu Thr Cys Pro Tyr Leu Lys Phe Ala His Phe Thr Ala
121 245 250 255
123 Asn Gln Ala Ile Leu Glu Ala Phe Gln Gly Lys Lys Arg Val His Val
124 260 265 270
126 Ile Asp Phe Ser Met Ser Gln Gly Leu Gln Trp Pro Ala Leu Met Gln
127 275 280 285
129 Ala Leu Ala Leu Arg Pro Gly Gly Pro Pro Val Phe Arg Leu Thr Gly

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/911,513

DATE: 08/01/2001

TIME: 17:10:22

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\08012001\I911513.raw

```

130      290      295      300
132 Ile Gly Pro Pro Ala Pro Asp Asn Phe Asp Tyr Leu His Glu Val Gly
133 305      310      315      320
135 Cys Lys Leu Ala His Leu Ala Glu Ala Ile His Val Glu Phe Glu Tyr
136      325      330      335
138 Arg Gly Phe Val Ala Asn Thr Leu Ala Asp Leu Asp Ala Ser Met Leu
139      340      345      350
141 Glu Leu Arg Pro Ser Glu Ile Glu Ser Val Ala Val Asn Ser Val Phe
142      355      360      365
144 Glu Leu His Lys Leu Leu Gly Arg Pro Gly Ala Ile Asp Lys Val Leu
145      370      375      380
147 Gly Val Val Asn Gln Ile Lys Pro Glu Ile Phe Thr Val Val Glu Gln
148 385      390      395      400
150 Glu Ser Asn His Asn Ser Pro Ile Phe Leu Asp Arg Phe Thr Glu Ser
151      405      410      415
153 Leu His Tyr Tyr Ser Thr Leu Phe Asp Ser Leu Glu Gly Val Pro Ser
154      420      425      430
156 Gly Gln Asp Lys Val Met Ser Glu Val Tyr Leu Gly Lys Gln Ile Cys
157      435      440      445
159 Asn Val Val Ala Cys Asp Gly Pro Asp Arg Val Glu Arg His Glu Thr
160      450      455      460
162 Leu Ser Gln Trp Arg Asn Arg Phe Gly Ser Ala Gly Phe Ala Ala Ala
163 465      470      475      480
165 His Ile Gly Ser Asn Ala Phe Lys Gln Ala Ser Met Leu Leu Ala Leu
166      485      490      495
168 Phe Asn Gly Gly Glu Gly Tyr Arg Val Glu Glu Ser Asp Gly Cys Leu
169      500      505      510
171 Met Leu Gly Trp His Thr Arg Pro Leu Ile Ala Thr Ser Ala Trp Lys
172      515      520      525
174 Leu Ser Thr Asn
175      530
178 <210> SEQ ID NO: 3
179 <211> LENGTH: 1643
180 <212> TYPE: DNA
181 <213> ORGANISM: Arabidopsis thaliana
183 <400> SEQUENCE: 3
184 tagaagtggg agtggagtga aaaaacaaat cctaagcagt cctaaccgat ccccgaaagt 60
185 aaagattctt caccttccca aataaagcaa aacctagatc cgacattgaa ggaaaaacct 120
186 tttagatcca tctctgaaaa aaaaccaacc atgaagagag atcatcatca tcatcatcaa 180
187 gataagaaga ctatgatgat gaatgaagaa gacgacggta acggcatgga tgttgctcag 240
188 aaactcgagc agcttgaagt tatgatgtct aatgttcaag aagacgatct ttctcaactc 300
189 gctactgaga ctgttcacta taatccggcg gagctttaca cgtggcttga ttctatgctc 360
190 accgacctta atcctccgtc gtctaaccgc gactacgata ttaaagctat tcccgggtgac 420
191 gcgattctca atcagttcgc tatcgattcg gcttcttcgt ctaaccaagg cggcgaggga 480
192 gatacgtata ctacaaacaa gcggttgaaa tgctcaaacy gcgtcgtgga aaccaccaca 540
193 gcgacggctg agtcaactcg gcatgttgtc ctggttgact cgcaggagaa cgggtgtcgt 600
194 ctcgttcacg cgtttttggc ttgcgctgaa gctgttcaga aggagaatct gactgtggcg 660
195 gaagctctgg tgaagcaaat cggattotta gctgtttctc aaatcggagc tatgagaaaa 720
196 gtcgctactt acttcgccga agctctcgcg cggcggattt accgtctctc tccgtcgcag 780

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/911,513

TIME: 17:10:22

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\08012001\I911513.raw

```

197 agtccaatcg accactctct ctccgatact ctttagatgc acttctacga gacttgtcct 840
198 tatctcaagt tcgctcactt cacggcgaat caagcgattc tcgaagcttt tcaagggag 900
199 aaaagagttc atgtcattga tttctctatg agtcaaggtc ttcaatggcc ggcgcttatg 960
200 caggctcttg cgcttcgacc tgggtggtcct cctgttttcc ggtaaccgg aattgggtcca 1020
201 ccggcaccgg ataatttcga ttatcttcat gaagttgggt gtaagctggc tcatttagct 1080
202 gaggcgattc acgttgagtt tgagtacaga ggatttgtgg ctaacacttt agctgatctt 1140
203 gatgcttcga tgcttgagct tagaccaagt gagattgaat ctgttgcggt taactctgtt 1200
204 ttcgagcttc acaagctctt gggacgacct ggtgcgatcg ataaggttct tgggtgtggtg 1260
205 aatcagatta aaccggagat ttctactgtg gttgagcagg aatcgaacca taatagtccg 1320
206 attttcttag atcgggtttac tgagtcggtt cattattact cgacgttggt tgactcgttg 1380
207 gaaggtgtac cgagtgttca agacaaggtc atgtcggagg tttacttggg taaacagatc 1440
208 tgcaacgttg tggttgtga tggaacctgac cgagttgagc gtcataaaac gttgagtcag 1500
209 tggaggaacc ggttcgggtc tgctgggttt gcgctgcac atattggttc gaatgcgttt 1560
210 aagcaagcga gtatgctttt ggctctgttc aacggcggtg agggttatcg ggtggaggag 1620
211 agtgacggct gtctcatggt ggg                                     1643
213 <210> SEQ ID NO: 4
214 <211> LENGTH: 221
215 <212> TYPE: PRT
216 <213> ORGANISM: Arabidopsis thaliana
218 <400> SEQUENCE: 4
219 Met Lys Arg Asp His His His His His Gln Asp Lys Lys Thr Met Met
220 1 5 10 15
222 Met Asn Glu Glu Asp Asp Gly Asn Gly Met Asp Val Ala Gln Lys Leu
223 20 25 30
225 Glu Gln Leu Glu Val Met Met Ser Asn Val Gln Glu Asp Asp Leu Ser
226 35 40 45
228 Gln Leu Ala Thr Glu Thr Val His Tyr Asn Pro Ala Glu Leu Tyr Thr
229 50 55 60
231 Trp Leu Asp Ser Met Leu Thr Asp Leu Asn Pro Pro Ser Ser Asn Ala
232 65 70 75 80
234 Glu Tyr Asp Leu Lys Ala Ile Pro Gly Asp Ala Ile Leu Asn Gln Phe
235 85 90 95
237 Ala Ile Asp Ser Ala Ser Ser Ser Asn Gln Gly Gly Gly Gly Asp Thr
238 100 105 110
240 Tyr Thr Thr Asn Lys Arg Leu Lys Cys Ser Asn Gly Val Val Glu Thr
241 115 120 125
243 Thr Thr Ala Thr Ala Glu Ser Thr Arg His Val Val Leu Val Asp Ser
244 130 135 140
246 Gln Glu Asn Gly Val Arg Leu Val His Ala Leu Leu Ala Cys Ala Glu
247 145 150 155 160
249 Ala Val Gln Lys Glu Asn Leu Thr Val Ala Glu Ala Leu Val Lys Gln
250 165 170 175
252 Ile Gly Phe Leu Ala Val Ser Gln Ile Gly Ala Met Arg Lys Val Ala
253 180 185 190
255 Thr Tyr Phe Ala Glu Ala Leu Ala Arg Arg Ile Tyr Arg Leu Ser Pro
256 195 200 205
258 Ser Gln Ser Pro Ile Asp His Ser Leu Ser Asp Thr Leu
259 210 215 220
262 <210> SEQ ID NO: 5

```

RAW SEQUENCE LISTING

DATE: 08/01/2001

PATENT APPLICATION: US/09/911,513

TIME: 17:10:22

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\08012001\I911513.raw

```

263 <211> LENGTH: 1642
264 <212> TYPE: DNA
265 <213> ORGANISM: Arabidopsis thaliana
267 <400> SEQUENCE: 5
268 tagaagtggg agtggagtgga aaaaacaaat cctaagcagt cctaaccgat ccccgagagct 60
269 aaagattctt caccttccca aataaagcaa aacctagatc cgacattgaa ggaaaaaacct 120
270 tttagatcca tctctgaaaa aaaaccaacc atgaagagag atcatcatca tcatcatcaa 180
271 gataagaaga ctatgatgat gaatgaagaa gacgacggta acggcatgga tgttgctcag 240
272 aaactcgagc agcttgaaagt tatgatgtct aatgttcaag aagacgatct ttctcaactc 300
273 gctactgaga ctgttcacta taatccggcg gagctttaca cgtggcttga ttctatgctc 360
274 accgacctta atcctccgct gtctaacgcc gagtacgata ttaaagctat tcccggtgac 420
275 gcgattctca atcagttcgc tatcgattcg gcttcttcgt ctaaccaagg cggcggagga 480
276 gatacgtata ctacaaacaa gcggttgaaa tgctcaaacg gcgtcgtgga aaccaccaca 540
277 gcgacggctg agtcaactcg gcatgttgct ctggttgact cgcaggagaa cgggtgtcgt 600
278 ctgcttcacg cgcttttggc ttgcgctgaa gctgttcaga aggagaatct gactgtggcg 660
279 gaagctctgg tgaagcaaat cggattctta gctgtttctc aaatcggagc tatgagaaaa 720
280 gtcgctactt acttcgccga agctctcgcg cggcggagatt accgtctctc tccgctcgag 780
281 agtccaatcg accactctct ctccgatact cttcagatgc acttctacga gacttgctct 840
282 tatctcaagt tcgctcactt cacggcgaat caagcgattc tcgaagcttt tcaagggaag 900
283 aaaagagttc atgtcattga ttctctatga gtcaaggctt tcaatggccg gcgcttatgc 960
284 aggtctttgc gcttcgacct ggtggtcctc ctgttttccg gtttaaccgga attggtccac 1020
285 cggcaccgga taatttcgat tatcttcatg aagttgggtg taagctggct catttagctg 1080
286 aggcgattca cgttgagttt gagtacagag gatttggtgc taacacttta gctgatcttg 1140
287 atgcttcgat gcttgagctt agaccaagtg agattgaatc tgttgcggtt aactctgttt 1200
288 tcgagcttca caagctcttg ggacgacctg gtgcgatacg taaggttctt ggtgtggtga 1260
289 atcagattaa accggagatt ttcaactgtg ttgagcagga atcgaaccat aatagtccga 1320
290 ttttcttaga tcggtttact gagtcgttgc attattaact gacgttggtt gactcgttgg 1380
291 aaggtgtacc gagtgggtcaa gacaagggtc tgctcgaggt ttacttggtt aaacagatct 1440
292 gcaacgttgt ggcttgatg ggacctgacc gagttgagcg tcatgaaacg ttgagtcagt 1500
293 ggaggaaccg gttcgggtct gctgggtttg cggctgcaca tattggttcg aatgcgttta 1560
294 agcaagcgag tatgcttttg gctctgttca acggcgggtg gggttatcgg gtggaggaga 1620
295 gtgacggctg tctcatgttg gg
297 <210> SEQ ID NO: 6
298 <211> LENGTH: 259
299 <212> TYPE: PRT
300 <213> ORGANISM: Arabidopsis thaliana
302 <400> SEQUENCE: 6
303 Met Lys Arg Asp His His His His His Gln Asp Lys Lys Thr Met Met
304 1 5 10 15
306 Met Asn Glu Glu Asp Asp Gly Asn Gly Met Asp Val Ala Gln Lys Leu
307 20 25 30
309 Glu Gln Leu Glu Val Met Met Ser Asn Val Gln Glu Asp Asp Leu Ser
310 35 40 45
312 Gln Leu Ala Thr Glu Thr Val His Tyr Asn Pro Ala Glu Leu Tyr Thr
313 50 55 60
315 Trp Leu Asp Ser Met Leu Thr Asp Leu Asn Pro Pro Ser Ser Asn Ala
316 65 70 75 80
318 Glu Tyr Asp Leu Lys Ala Ile Pro Gly Asp Ala Ile Leu Asn Gln Phe
319 85 90 95

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/911,513

DATE: 08/01/2001

TIME: 17:10:23

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\08012001\I911513.raw

L:13 M:270 C: Current Application Number differs, Replaced Application Number